

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/KR05/000526

International filing date: 26 February 2005 (26.02.2005)

Document type: Certified copy of priority document

Document details: Country/Office: KR
Number: 10-2004-0013016
Filing date: 26 February 2004 (26.02.2004)

Date of receipt at the International Bureau: 17 May 2005 (17.05.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto
is a true copy from the records of the Korean Intellectual
Property Office

출 원 번 호 : 특허출원 2004년 제 0013016 호
Application Number 10-2004-0013016

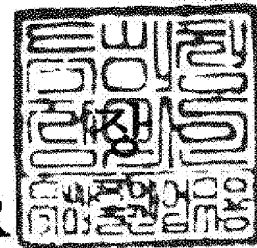
출 원 일 자 : 2004년 02월 26일
Date of Application FEB 26, 2004

출 원 인 : 남승현
Applicant(s) Nam, Seung Hyon

2005 년 04 월 07 일

특 허 청

COMMISSIONER



【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【제출일자】	2004.02.26
【국제특허분류】	G06F
【국제특허분류】	G10L
【국제특허분류】	H03H
【발명의 국문명칭】	다중경로 다채널 혼합신호의 주파수 영역 블라인드 분리를 위한 방법 및 그 장치
【발명의 영문명칭】	The Methods and Apparatus for Blind Separation of Multichannel Convolutive Mixtures in the Frequency-domain
【출원인】	
【성명】	남승현
【출원인코드】	4-2004-000792-1
【지분】	100/100
【발명자】	
【성명】	남승현
【출원인코드】	4-2004-000792-1
【심사청구】	청구
【취지】	특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정 에 의한 출원심사를 청구합니다. 출원인 남승현 (인)
【수수료】	
【기본출원료】	32 면 38,000 원
【가산출원료】	0 면 0 원

【우선권주장료】	0 건 0 원
【심사청구료】	7 항 333,000 원
【합계】	371,000 원
【감면사유】	개인(70%감면)
【감면후 수수료】	111,300 원
【첨부서류】	1.요약서·명세서(도면)_1통

【요약서】

【요약】

1. 청구범위에 기재된 발명이 속한 기술 분야

본 발명은 다중 경로에 의하여 혼합된 독립적인 신호들을 신호에 대한 사전 정보 없이 분리하는 다중경로 다채널 혼합신호의 주파수 영역 블라인드 분리(Blind Source Separation)를 위한 방법 및 그 장치와 상기 방법을 실현시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체에 관한 것임

2. 발명이 해결하고자 하는 기술적 과제

본 발명은 실제 소음 환경에서 다중 경로를 통해 혼합되어 여러 개의 센서(Sensor)를 통해 수집된 음성 및 오디오 신호들로부터 분리필터를 추정하여 원신호를 분리해내는 경우에, 분리필터의 빠른 수렴을 위해 적용되는 자연접선기울기를 주파수 영역에서 효과적으로 적용시킴으로써, 빠르고 안정적인 적응 분리회로를 구현할 수 있는 주파수 영역의 다중경로 다채널 블라인드 혼합신호 분리를 위한 장치 및 그 방법을 제공함.

3. 발명의 해결 방법의 요지

본 발명의 주파수 영역의 다중경로 다채널 블라인드 혼합신호 분리를 위한 장치는, 수집된 다채널 혼합신호를 주파수 영역에서 블록(Block) 단위로 분리하기 위해, 수집된 다채널 혼합신호로부터 여러 개의 블록들로 구성된 현재 프레임(Frame)을 형성하는 수단; 상기 현재 프레임의 신호를 주파수 영역에서 분리하는

수단; 상기 분리된 신호를 비선형함수를 통해 변환하는 수단; 상기 분리신호와 비선형 변환된 신호를 이용하여 정규화된 상호전력 스펙트럼(Cross-Power Spectrum)을 계산하는 수단; 상기 정규화된 상호전력 스펙트럼을 이용하여 자연접선 기울기를 계산하는 수단; 상기 자연접선 기울기를 이용하여 분리필터 계수를 갱신하는 수단; 상기 갱신된 분리필터계수를 정규화하는 수단; 상기 분리필터의 수렴을 점검하는 수단; 상기 분리필터가 수렴된 경우 분리필터를 이용하여 신호를 분리하는 수단을 포함함.

4. 발명의 중요한 용도

블라인드 신호분리 기술은 다중 마이크로폰 (Microphone) 어레이를 이용한 화자 추적, 열화된 신호의 음성인식이나 음질 향상을 위한 음성부호화기의 전 처리기, 3차원 객체지향 오디오 신호처리를 위한 신호 분리기, 스테레오 반향제거기 (Stereophonic Echo Canceller), 통신 채널의 간섭제거, 도래각(Direction of Arrival) 추정이나 EEG나 MEG과 같은 생체신호 등에서 광범위하게 활용됨.

【대표도】

도 3

【색인어】

블라인드 신호 분리, 다채널 블라인드 디콘볼루션(Multichannel Blind

Deconvolution), 자연접선기울기, 논홀로노믹(Nonholonomic)

【명세서】

【발명의 명칭】

다중경로 다채널 혼합신호의 주파수 영역 블라인드 분리를 위한 방법 및 그 장치{The Methods and Apparatus for Blind Separation of Multichannel Convulsive Mixtures in the Frequency-domain}

【도면의 간단한 설명】

- <1> 도 1은 다채널 다중경로 혼합신호 블라인드 분리 장치에 대한 구성도.
- <2> 도 2a는 자연접선기울기를 이용한 다채널 블라인드 디콘볼루션 알고리즘의 블록도.
- <3> 도 2b는 다채널 혼합신호를 분리하기 위한 다채널 분리필터 다이어그램 (2채널 예).
- <4> 도 2c는 메모리가 없는 비선형 함수를 이용한 혼합신호의 확률밀도를 균일한 확률밀도로 변환하는 개념도.
- <5> 도 3은 자연접선기울기를 이용한 주파수 영역 다채널 블라인드 디콘볼루션 알고리즘의 순서도.
- <6> 도 4a는 실제 사무실 환경에서 녹음된 혼합된 음성-음악신호(x_1, x_2)와 본 발명의 방법을 이용하여 분리한 음성-음악 분리신호(u_1, u_2).
- <7> 도 4b는 실제 사무실 환경에서 녹음된 혼합된 음성-음악신호를 본 발명의 분

리방법을 이용하여 분리한 결과 최종적으로 수렴된 분리필터 계수.

<8> 도 5는 두 사람의 원음성(s_1, s_2), 가상의 혼합회로를 통해 혼합된 음성(x_1, x_2), 본 발명의 분리 방법을 분리한 음성(u_1, u_2)의 파형 비교.

<9> <도면의 주요 부분에 대한 부호의 설명>

<10> 100 : 분리 장치 102 : 센서 어레이

<11> 104 : 신호처리 프로세서 106 : 녹음장치

<12> 108 : 컴퓨터 시스템 110 : 입력 장치

<13> 112 : 출력 장치 114 : 중앙처리장치

<14> 116 : 보조회로 118 : 음성인식기/음성부호화기

<15> 120 : 입출력 인터페이스 122 : 메모리

<16> 124 : 주파수 영역 다채널 블라인드 디콘볼루션 프로그램

<17> 126 : 다채널 혼합신호 입력 장치

<18> 201 : 분리부 202 : 비선형 변환부

<19> 203 : 자연접선기울기를 이용한 분리필터 계수 갱신 연산부

<20> 301 : 다채널 혼합신호 입력 유닛 302 : 현재 프레임 형성 유닛

<21> 303 : 현재 프레임 신호 분리 유닛 304 : 비선형 함수 변환 유닛

<22> 305 : 정규화된 상호전력 스펙트럼 계산 유닛

<23> 306 : 상호전력 스펙트럼 계산 유닛 307 : 평균전력 스펙트럼 유닛

<24> 308 : 상호전력 스펙트럼 정규화 유닛

- <25> 309 : 시간 영역의 제약조건 적용 유닛 310 : 자연접선기울기 계산 유닛
- <26> 311 : 분리필터 계수 갱신 유닛
- <27> 312 : 분리필터 계수 정규화 유닛
- <28> 313 : 수렴여부 판단 유닛 314 : 신호분리 출력 유닛

【발명의 상세한 설명】

【발명의 목적】

【발명이 속하는 기술분야 및 그 분야의 종래기술】

<29> 본 발명은 신호처리 분야에 속하며 보다 구체적으로는 다중경로 다채널 혼합 신호의 주파수 영역 블라인드 분리를 위한 방법 및 그 장치와 상기 방법을 실현하기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체에 관한 것이다.

<30> 실제 소음환경에서 음성이나 오디오 신호들은 다중 경로를 통해 혼합되어 여러 센서들에 인가되므로 원 신호를 얻기 위해서는 혼합된 신호들을 분리할 필요가 있다. 이때 원 신호에 대한 사전 정보 없이 신호를 분리하는 기술을 블라인드 신호 분리 (Blind Source Separation; BSS) 기술이라고 부른다. 블라인드 신호분리는 다중 화자 혹은 소나 어레이(Sonar Array) 등과 같은 서로 독립적인 신호들이 다중 경로를 통해 혼합되었을 때 이를 분리하는데 매우 효과적인 기술이다. 블라인드 신호분리 기술은 다중 마이크로폰 (Microphone) 어레이를 이용한 화자 추적, 열화된 신호의 음성인식이나 음질 향상을 위한 음성부호화기의 전 처리기, 3차원 객체지향

오디오 신호처리를 위한 신호 분리기, 스테레오 반향제거기(Stereophonic Echo Canceller), 통신 채널의 간섭제거, 도래각(Direction of Arrival) 추정이나 EEG나 MEG과 같은 생체신호 등에서 광범위하게 활용된다.

<31> 대부분의 블라인드 신호 분리 방식은 다중 혼합 경로의 영향을 무효화함으로써 원 신호를 복원하고자 한다. 다중 혼합 경로의 영향을 무효화하는데 필요한 이상적인 다중 분리에는 무한 길이의 필터가 필요하지만 대부분의 실용 가능한 상황에서는 유한 길이의 분리필터로서도 충분히 가능하다. 특히, 화자나 물체의 이동이나 환경의 변화에 충분히 빠르게 적응적으로 반응하기 위한 블라인드 신호분리 방법이 필요하다.

<32> 현재 널리 사용되고 있는 다채널 블라인드 신호 분리의 두 가지 유형은 (1) 독립적인 신호들의 2차 통계 특성을 이용한 다중 상관관계 제거 (Multiple Decorrelation) 방식과 (2) 비선형 함수를 이용한 다채널 블라인드 디콘볼루션 (Multichannel Blind Deconvolution: MBD) 방식이다.

<33> 대표적인 다중 상관관계 제거 방식으로는 최근 제안된 Convolutional Blind Source Separation Using A Multiple Decorrelation Method (Lucas Parra and Clay Spence, US Patent 6167417)가 있다. 이 방법은 수렴 속도가 빠르고 안정적인 특징이 있지만 주파수 영역에서 분리신호들이 서로 뒤섞이는 (Frequency Permutation) 단점이 있다. 또한 대각 분리필터를 1로 고정하고 비대각 분리필터만을 가지고 분리하기 때문에 분리 성능이 떨어지는 단점도 있다.

<34> MBD 방식에는 여러 방법들이 제안되었지만 그중 다중경로 혼합에서 가장 좋

은 성능을 나타내는 것으로는 자연접선기울기를 이용한 MBD 방식이다. 이 방식은 원 신호가 평탄한 스펙트럼을 갖는 통신채널의 등화에는 매우 성공적이지만 음성이나 오디오 신호와 같이 스펙트럼이 기울어진 신호에 대하여는 매우 부적절하다. 그 이유는 MBD 방식이 평탄한 스펙트럼을 가정하여 유도되었기 때문에 분리 스펙트럼이 평탄해질 때까지 분리과정이 진행되기 때문이다. 이러한 현상을 백색화 현상(Whitening Effect)이라고 부른다.

<35> 최근 이러한 문제점을 극복하기 위해 주파수 영역에서 자연접선기울기를 사용하는 MBD 방식의 블라인드 신호분리가 고안되었다 (다중경로 혼합신호 분리 장치 및 그 방법, 특허출원번호:10-2003-36759,출원일:2003-6-9). 그러나 이 방식의 경우 자유접선기울기가 여전히 논홀로노믹(Nonholonomic) 특성을 충분히 만족시켜주지 못하므로 분리된 신호의 음질이 크게 저하되며 안정성이 떨어지는 단점이 있다. 본 발명에서는 자연접선기울기가 주파수 영역에서 논홀로노믹 특성을 만족하도록 정규화하는 방법을 제공함으로써 수렴특성을 안정화시키고, 분리성능을 크게 향상시키고, 백색화 현상을 효율적으로 제거함으로써 음질을 크게 향상시킨다.

【발명이 이루고자 하는 기술적 과제】

<36> 상기의 단점들은 본 발명에서 제안하는 에일리어싱이 제거된 정규화된 상호 전력 스펙트럼을 사용하는 블록 주파수 영역의 블라인드 신호분리 방법과 그 장치를 통해 극복된다. 구체적으로, 본 발명은 블라인드 신호 분리 방식이 잘못된 필터로 수렴하는 것을 방지하기 위해 상호전력 스펙트럼을 계산할 때 과거 프레임으로

부터의 분리신호를 사용하지 않고 오직 현재 프레임에서 분리된 신호만을 사용한다. 또한 에일리어싱이 제거된 정규화된 상호전력 스펙트럼을 사용함으로써 빠르고 안정적인 수렴을 보장하는 동시에 우수한 분리 성능과 우수한 분리신호 음질을 제공한다. 우수한 분리신호 음질은 음성인식기나 음성부호화기와 같은 응용에서는 매우 중요한 요소이다.

<37>

【발명의 구성】

<38>

본 발명은 블라인드 신호 분리를 위한 다중 경로 분리회로 필터를 추정하기 위하여 주파수 영역에서 안정적인 정규화된 자연접선기울기를 이용하는 다채널 블라인드 신호분리 방법과 그 장치를 제공한다.

<39>

도 1은 본 발명의 온라인(On-Line)과 오프라인(Off-Line) 신호 분리 방법을 구현하기 위한 시스템(100)이다. 시스템은 다중경로 혼합신호를 제공하는 입력 장치(126)와 이를 분리하는 주파수 영역 다채널 블라인드 디콘볼루션 프로그램(124)이 내장된 컴퓨터 시스템(108)으로 구성되어 있다. 다중경로 혼합신호 입력 장치(126)은 센서 어레이(102)와 다중경로 혼합신호를 디지털 신호로 변환하는 신호처리 프로세서(104)으로 구성되어 있으며 센서 어레이(102)는 마이크로폰과 같은 하나 이상의 센서들(102A, 102B, 102C)로 구성되어 있다.

<40>

컴퓨터 시스템(108)은 중앙처리장치(114), 메모리(122), 입출력 인터페이스(120), 보조회로(116)으로 구성되어 있다. 컴퓨터 시스템(108)은 마우스(Mouse), 키보드(Keyboard)등과 같은 입력장치(110)와 모니터(Monitor)나 프린터(Printer)와

같은 출력장치(112)와 연결된다. 보조회로(116)는 전원공급장치, 캐쉬(Cache), 타이밍회로(Timing Circuit), 통신회로(Communication Circuit), 버스(Bus)등과 같은 장치들이다. 메모리는 RAM(Random Access Memory), ROM(Read Only Memory), 디스크 드라이브(Disk Drive), 테이프 드라이브(Tape Drive), 플래쉬 메모리(Flash Memory), CD(Compact Disk) 혹은 이들의 조합과 같은 장치들이다. 본 발명은 주파수 영역 다중채널 블라인드 디콘볼루션 프로그램(124)로 구현되어 메모리(122)에 저장되어 있으며 중앙처리장치(114)에서 수행되어 다중경로 다채널 혼합신호 입력장치를 통하여 입력된 혼합신호(126)를 분리한다. 이러한 컴퓨터 시스템(108)은 본 발명을 실행할 수 있는 일반적인 컴퓨터 시스템이지만, 특별하게 고안된 ASIC(Application Specific Intergrated Circuits)전용이나 디지털 신호처리 프로세서(Digital Signal Processor: DSP)와 같은 컴퓨터 시스템을 포함한다. 이와 같이, 본 발명은 하드웨어(Hardware)나 소프트웨어(Software) 혹은 이들의 조합형태로 구현될 수 있다.

<41> 본 발명에서 컴퓨터 시스템(108)의 동작을 설명하기 위한 일 실시 예로 음성 인식기나 음성부호화기 등과 같은 신호처리 모듈(118)을 이용한다. 여러 사람이 함께 있는 회의실에서 사람들의 음성 혹은 음악과 배경 잡음들은 마이크로폰 어레이(102)를 통해 채집된다. 마이크로폰(102)에 의해 채집된 음성신호들은 다중경로를 통해 혼합된 다채널 신호들로서 만약 각 사람의 음성을 분리하여 각 음성을 음성부호화기에 인가할 수 있다. 혼합된 신호는 신호처리 프로세서(104)를 통해 필터링(Filtering)되고, 증폭된 다음 디지털로 변환되어 컴퓨터 시스템(108)로 보내진다.

중앙처리장치(114)는 주파수영역 다채널 블라인드 디콘볼루션 프로그램(124)를 수행함으로써 혼합신호를 각각의 독립적인 신호로 분리한다. 이 과정에서 혼합신호에 포함되었던 배경 잡음은 제거되고, 잡음이 제거된 신호는 음성인식기(혹은 음성부호화기)(118)에 보내진다. 음성인식기가 사용된 경우라면 음성은 음성인식기(118)에서 이를 해석하여 컴퓨터 명령이나 문자로 변환할 수 있으며, 음성부호화기가 사용된 경우라면 보다 깨끗한 음질의 통화를 제공한다. 이처럼, 컴퓨터 시스템(108)은 주파수 영역 다채널 블라인드 디콘볼루션 프로그램을 수행함으로써 음성인식기나 음성부호화기에 전 처리기로서 활용될 수 있다.

<42>

도 2a는 본 발명 주파수 영역 다채널 블라인드 디콘볼루션 프로그램(124)의 일실시에 구조도이다. 본 발명의 다중경로 다채널 분리 장치는 분리부(201), 비선형 변환부(202) 및 자연접선기울기를 이용한 분리필터계수 갱신 연산부(203)를 포함하고 있다. 상기 분리부(201)는 다중 채널 혼합신호 $\mathbf{x}(k)$ 로부터 신호를 분리하는 기능을 담당한다. 상기 혼합신호 $\mathbf{x}(k)$ 는 m 개의 독립적인 원 신호가 다중 경로를 통해 n 개의 센서에 인가된 것으로 다음과 같이 정의된다.

【수학식 1】

$$\mathbf{x}(k)=[x_1(k),x_2(k),\cdots,x_n(k)]^T$$

<44>

여기서, $x_j(k)$ 는 j 번째 센서에서 인가된 혼합신호이다. 이를 분리하기 위

한 분리필터는 $m \times n$ 행렬로 (i,j) 필터 원소를 z 변환의 형태로 표현하면 다음과 같다.

【수학식 2】

$$w_{ij}(z,k) = \sum_{p=0}^{L-1} w_{ij,p}(k) z^{-p}$$

<46> 이로부터, 분리된 신호 $\mathbf{u}(k)$ 는 다음과 같이 정의된다.

【수학식 3】

<47>
$$\mathbf{u}(k) = [u_1(k), u_2(k), \dots, u_m(k)]^T$$

<48> 여기서, $u_i(k)$ 는 i 번째 분리신호로서 다음과 같이 정의된다.

【수학식 4】

$$u_i(k) = \sum_{j=1}^n w_{ij,p}(k) x_j(k-p), \quad i=1, \dots, m.$$

<50> 도 2b는 $n=m=2$ 인 경우에 대한 분리필터 다이어그램이다. 상기 분리부

(201)에서 분리된 신호 $\mathbf{u}(k)$ 는 비선형 변환부(202)로 입력된다.

<51> 비선형 변환부(202)는 분리신호를 버스강(Bussgang) 유형의 메모리가 없는 비선형함수를 통해 변환하는 부분으로 다음과 같이 정의된다.

【수학식 5】

$$y_i(k)=g(u_i(k))$$

<53> 비선형 함수 $g(\cdot)$ 는 도 2c에서 볼 수 있는 바와 같이 원 신호의 확률밀도 $p(u)$ 를 균일한 확률밀도 $p(y)$ 로 변환한다. 비선형 함수는 원 신호의 확률밀도에 따라 다르게 선정된다. 음성이나 오디오 신호의 확률밀도는 감마(Gamma) 혹은 라플라시안(Laplacian)이며 이 경우 $\alpha \sin(u)$ 혹은 $\tanh(u)$ 와 같은 비선형 함수가 적절하다.

<54> 자연접선기울기를 이용한 분리필터 계수 갱신 연산부(203)는 다음과 같이 최대기울기 방식을 이용하여 분리필터 계수를 계산한다.

【수학식 6】

$$w_{ij,p}(k+1)=w_{ij,p}(k)+\mu \Delta w_{ij,p}(k), 1 \leq i \leq m, 1 \leq j \leq n, 0 \leq p \leq L-1$$

<56> 여기서 $\Delta w_{ij,p}$ 는 자연접선기울기로서 다음과 같이 주어진다

【수학식 7】

$$\Delta w_{ij,p}(k)=w_{ij,p}(k)-\sum_{k=1}^m \sum_{q=0}^p y_i(k) u_k(k-p+q) w_{ij,q}(k)$$

<58> 주목할 것은 (7)의 합에서 q 의 범위가 기존의 자연접선기울기를 이용한 MBD 알고리즘의 경우와는 달리 $(L-1)$ 이 아닌 p 로 제한된다는 점이다. 기존의 MBD 알

고리즘은 양방향 분리필터를 가정하므로 미래의 분리신호 샘플 값들을 필요로 한다. 기존의 자연접선기울기를 이용한 MBD 알고리즘에서는 이러한 비인과적인 연산을 피하기 위해 분리신호를 L 샘플만큼 지연시킨 다음 사용하는 지연 갱신을 적용한다. 이 지연갱신은 수렴속도가 저하되며 수렴특성의 저하를 초래하는 원인이 되었다. 이에 비해, 본 발명에서는 인과성의 최소위상을 갖는 단일방향 분리필터를 가정하기 때문에 미래의 분리신호 값을 필요로 하지 않았으며 결과적으로 지연의 필요성이 제거됨으로 인해 수렴속도와 수렴특성을 개선할 수 있다. 비 인과성 특성을 갖는 양방향 분리필터를 사용하는 근본적인 이유는 이상적인 분리필터가 비 최소위상을 갖는다는 점 때문이다. 현실적으로 응용 가능한 많은 현실문제에서 인과성 분리필터만으로도 우수한 분리성과를 얻을 수 있다는 사실은 잘 알려져 있다.

<59> 본 발명에서는 이러한 과정을 효율적으로 수행하기 위하여 DFT(Discrete Fourier Transform)을 이용하여 주파수 영역에서 블록 단위로 수행하는 오버랩-세이브(Overlap-Save)방식을 취한다. 여기서, 블록 크기 M , 프레임 크기 N , 필터 길이 L 로 정의하며 프레임 크기는 오버랩-세이브 방식에서 사용하는 오버랩의 정도에 따라 $M=N/r$ 에 의해 결정된다. 일반적으로 $r \geq 2$ 이며 50% 오버랩을 사용하는 경우 $r=2$ 로 설정하지만 빠른 수렴을 위해 2보다 큰 r 값을 사용할 수 있다. 여기서는 $r=2$ 로 가정된다. 또한 DFT와 IDFT의 크기는 $N=2^l$ 로 가정된다. 도 3

은 본 발명의 다채널 다중경로 혼합신호의 주파수 영역 분리방법에 대한 순서도이다.

<60> 단계 301의 다채널 혼합신호 $\mathbf{x}(k)$ 는 식 (1)과 동일하다.

<61> 단계 302에서는 블록 시간 b 일 때 다채널 혼합신호를 각 채널별로 M 샘플의 블록으로 인가하여 과거 블록과 묶어서 N 샘플 단위의 프레임을 형성한다.

【수학식 8】

$$\mathbf{x}_j(b)=[x_j(bM-2M+1),\cdots,x_j(bM)]^T$$

<63> 단계 303는 블록시간 b 에서의 분리회로 계수

【수학식 9】

$$\mathbf{w}_{ij}(b)=[w_{ij,0},w_{ij,1},\cdots,w_{ij,L-1}]^T$$

<65> 를 이용하여 입력된 다채널 혼합신호를 분리하는 과정이다. 분리필터 계수는 만약 분리필터 계수에 대한 유용한 정보가 있다면 그 정보를 이용하여 적절하게 초기화될 수 있으며 일반적으로는 다음과 같이 초기화된다.

【수학식 10a】

$$\mathbf{w}_{ij}(b)=[1,0,\cdots,0]^T, i=j$$

【수학식 10b】

$$\mathbf{w}_{ij}(b) = [0, \dots, 0]^T, i \neq j$$

<68> 이로부터 분리신호는 주파수 영역에서 다음과 같이 계산된다.

【수학식 11】

$$\mathbf{u}_i(f, b) = \sum_{j=1}^n \mathbf{w}_{ij}(f, b) \odot \mathbf{x}_j(f, b)$$

<70> 여기서 \odot 는 원소별 곱셈이며 $\mathbf{w}_{ij}(f, b)$ 와 $\mathbf{x}_j(f, b)$ 는 $N \times 1$ 벡터들로서

각각 $\mathbf{w}_{ij}(b)$ 와 $\mathbf{x}_j(b)$ 의 푸리에 변환으로 다음과 같이 정의된다.

【수학식 12a】

$$\mathbf{w}_{ij}(f, b) = \mathbf{F} \mathbf{w}_{ij}(b)$$

【수학식 12b】

$$\mathbf{x}_j(f, b) = \mathbf{F} \mathbf{x}_j(b)$$

<73> 여기서 \mathbf{F} 는 $N \times N$ 크기의 N -point DFT 행렬을 의미하며 DFT는 일반적으로 FFT(Fast Fourier Transform)을 이용하여 계산된다. 만약 변환하고자 하는 벡터의 길이가 N 보다 작을 때 벡터의 후반부에 0을 채워서 길이 N 의 벡터로 만든다. 주파수 영역에서의 컨볼루션은 순환 컨볼루션(Circular Convolution)이며 선형 컨볼루션(Linear Convolution)을 계산하기 위해서는 시간 영역에서 에일리어싱이 일

어난 처음 $(L-1)$ 개의 샘플들을 버리고 후반부 $(N-L+1)$ 샘플만을 취한다. 일반적인 50% 오버랩-세이프 방식에서는 처음 M 샘플의 한 블록을 다 버리고 후반부의 M 샘플만을 취하지만, 본 발명에서는 초반 L 개의 샘플만을 버림으로써 다음에 전력 스펙트럼을 계산하는 과정에서 정확한 값을 계산하는 동시에 알고리즘을 안정하게 만든다. 에일리어싱을 제거하는 과정은 다음과 같이 정의된다.

【수학식 13】

$$\mathbf{u}_i(b) = \mathbf{P}_{0,N-L} \mathbf{F} \mathbf{u}_i(f,b) = [0, \dots, 0, u_i(bM-2M+L), \dots, u_i(bM)]^T$$

<75> 여기서, $\mathbf{P}_{0,N-L}$ 은 처음 L 개의 샘플값을 0으로 만드는 $N \times N$ 크기의 프로젝션(projection) 행렬로 다음과 같이 정의된다.

【수학식 14】

$$\mathbf{P}_{0,N-L} = \begin{bmatrix} \mathbf{0}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{N-L} \end{bmatrix}$$

<77> 여기서, $\mathbf{0}_L$ 은 $L \times L$ 크기의 0행렬이고 \mathbf{I}_{N-L} 은 $(N-L) \times (N-L)$ 크기의 단위행렬이다.

<78> 단계 304는 분리된 신호를 버스강 유형의 비선형함수를 통해 변환하는 과정으로 다음의 두 가지 방법 중 하나를 사용할 수 있다.

【수학식 15a】

$$\mathbf{y}_i(b) = g(\mathbf{u}_i(b)) = [0, \dots, 0, g(u_i(bM-2M+L)), \dots, g(u_i(bM))]^T$$

【수학식 15b】

$$\mathbf{y}_i(b) = g(\mathbf{u}_i(b)) = [0, \dots, 0, g(u_i(bM-2M+2L)), \dots, g(u_i(bM))]^T$$

<81> 일반적으로 식 (15a)를 사용하면 비선형함수와 분리신호와 상호상관관계 계수 (Cross-correlation)인 $g(u_i(k))u_j(k-p), p=0,1,\dots,L-1$, 에는 바이어스(bias)가 포함되며 식 (15b)를 사용하면 바이어스가 없는 계수를 얻게 된다.

<82> 단계 305는 에일리어싱이 제거된 정규화된 상호전력 스펙트럼을 계산하는 과정으로 본 발명에서 매우 중요한 과정으로 단계 306~309로 구성된다.

<83> 단계 306은 단계 303과 304에서 형성된 분리신호 $\mathbf{u}_j(b)$ 와 비선형 변환된 신호 $\mathbf{y}_i(b)$ 를 각각 DFT 변환하고 이를 이용하여 다음과 같이 블록 b 에서의 상호 전력 스펙트럼을 계산한다.

【수학식 16】

$$\mathbf{P}_{y_i u_j}(f, b) = \mathbf{y}_i(f, b) \odot \mathbf{u}_j^*(f, b)$$

<85> 여기서, $*$ 는 켈레복소수를 의미하며 $\mathbf{y}_i(f, b)$ 와 $\mathbf{u}_j(f, b)$ 는 $N \times 1$ 벡터들로서 각각 다음과 같이 정의된다.

【수학식 17a】

$$\mathbf{y}_i(f,b) = \mathbf{F} \mathbf{y}_i(b)$$

【수학식 17b】

$$\mathbf{u}_j(f,b) = \mathbf{F} \mathbf{u}_j(b)$$

<88> 여기서 주목할 것은 상호전력 스펙트럼의 계산에 사용되는 분리신호 $\mathbf{u}_j(b)$ 와 비선형 변환된 신호 $\mathbf{y}_i(b)$ 가 수식 (13), (15a), (15b)에 나타난 것처럼 현재 프레임에서 분리된 신호만을 이용하여 구성되었다는 점과 $\mathbf{u}_j(b)$ 의 처음 L 샘플들이 0이라는 점이다. 이것은 일반적으로 $\mathbf{u}_j(b)$ 가

【수학식 18】

$$\mathbf{u}_i(b) = [\mathbf{u}_i(bM-2M+1), \dots, \mathbf{u}_i(bM)]^T$$

<90> 다음과 같이 모두 채워진 형태로 조성된다는 점과 크게 다르다. 수학식 (18)에서 처음 L 샘플들은 $(b-1)$ 블록 시간에서의 분리필터 값을 이용하여 계산된 분리신호이며 나머지 $(N-2L)$ 샘플들은 b 블록 시간의 분리필터 계수를 이용하여 계산된 값이다. 따라서 분리신호열의 일관성이 파괴되어 상호전력 스펙트럼이 부정확하게 산출되며 이로 인해 알고리즘 수렴과정이 불안정해지고 분리성능이 저하된다. 뿐만 아니라, 알고리즘의 구현에서 과거 프레임(혹은 블록)의 분리신호 값

과 비선형함수 변환된 신호 값을 저장해야하는 부담이 발생한다.

<91>

단계 307은 상호전력 스펙트럼을 정규화하기 위해 분리신호와 비선형변환신호의 평균전력 스펙트럼들을 계산하는 과정이다. 일반적으로 신호들은 시간에 따라 많이 변화하므로 매 블록 마다 다음의 순환식을 이용하여 평균전력을 갱신한다.

【수학식 19a】

$$P_{y_i}(f,b) = (1-\gamma)P_{y_i}(f,b-1) + \gamma|y_i(f,b)|^2, i=1, \dots, m$$

【수학식 19b】

$$P_{u_j}(f,b) = (1-\gamma)P_{u_j}(f,b-1) + \gamma|u_j(f,b)|^2, j=1, \dots, m$$

<94>

여기서, γ 는 0과 1사이의 상수로서 신호의 유형과 환경변화에 따라 적절하게 선택될 수 있으며, 평균전력 스펙트럼은

$P_{y_i}(f,b) = P_{u_j}(f,0) = c[1, \dots, 1]^T, c < 1,$ 로 초기화된다. 단계 308은 상호전력 스펙트럼을 정규화하는 과정이다. 이전 발명에서는 상호전력스펙트럼을 정규화하기 위해 각 주파수에서 상호전력 스펙트럼의 모든 원소들을 동시에 각자의 평균전력 스펙트럼으로 정규화하였다. 본 발명에서는 정규화된 상호전력 스펙트럼이 각 주파수에서 단위행렬로 수렴하도록 만들어주는 동시에 비대각 원소들은 효과적으로 0으로 수렴하도록 만들어준다. 이로 인해, 신호분리의 필연적 결과인 분리신호간의 독립성을 얻을 뿐 아니라, 대각원소들이 효과적으로 정규화 됨으로써 알고리즘의 수렴특성을 크게 개선하였고 우수한 분리 성능을 얻을 수 있게 되었다. 뿐만 아니라,

최종적으로 분리된 분리신호에서 나타나는 백색화 현상(Whitening Effect)가 현저하게 줄어들었다.

<95> 이제, 정규화는 다음과 같이 수행된다.

【수학식 20】

$$\overline{P}_{y_{ij}}(f,b) = P_{y_{ij}}(f,b) \oslash \sqrt{P_{y_i}(f,b) \odot P_{y_j}(f,b)}$$

<97> 여기서, \oslash 는 원소별 나눗셈이다. 수학식 (20)의 정규화된 상호전력 스펙트럼을 시간 영역으로 IDFT 변환하면 상호상관관계 계수(Cross-Correlations)가 된다. 그러나 이 값들은 주파수 영역에서 순환 상관연산(Circular Correlation) (16)을 통해 계산된 것으로 에일리어싱 부분을 포함하고 있다. 따라서 에일리어싱 부분은 제거되어야만 한다. 뿐만 아니라, 단계 310의 자연접선기울기 계산에서 필요로 하는 부분은 분리회로의 길이와 동일한 처음 L 개의 상관관계 계수 값이므로 나머지 후반의 $(N-L)$ 값들은 0으로 대체되어야만 한다. 이러한 제한조건을 만족시키는 것은 순수한 최소위상의 특성을 가진 분리회로를 사용하기 때문이다. 단계 309에서 적용되는 시간 영역의 제한조건은 다음과 같이 정의된다.

【수학식 21】

$$\tilde{P}_{y_{ij}}(f,b) = F P_{L,0} F^{-1} \overline{P}_{y_{ij}}(f,b)$$

<99> 여기서 F^{-1} 는 $N \times N$ 크기의 N -point 역 DFT 행렬을 의미하며 $P_{L,0}$ 은

처음 L 개의 샘플은 그대로 남겨놓고 나머지 $(N-L)$ 개의 샘플 값을 0으로 만드는 $N \times N$ 크기의 프로젝션(projection) 행렬로 다음과 같이 정의된다.

【수학식 22】

$$\mathbf{P}_{L,0} = \begin{bmatrix} \mathbf{I}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{N-L} \end{bmatrix}$$

<101> 여기서, \mathbf{I}_L 은 $L \times L$ 크기의 단위행렬이고 \mathbf{I}_{N-L} 은 $(N-L) \times (N-L)$ 크기의 0 행렬이다.

<102> 단계 310는 논홀로노믹(nonholonomic) 조건을 적용하여 자연접선기울기를 계산하는 과정으로 다음과 같다.

【수학식 23a】

$$\hat{\mathbf{P}}_{y_{i\mathcal{U}_j}}(f,b) = \begin{cases} \bar{\mathbf{1}} - \tilde{\mathbf{P}}_{y_{i\mathcal{U}_j}}(f,b), & \text{for } i=j \\ -\tilde{\mathbf{P}}_{y_{i\mathcal{U}_j}}(f,b), & \text{for } i \neq j \end{cases}$$

【수학식 23b】

$$\Delta \mathbf{w}_{ij}(f,b) = \sum_{k=1}^m \hat{\mathbf{P}}_{y_{i\mathcal{U}_k}}(f,b) \mathbf{w}_{kj}(f,b)$$

<105> 여기서, $\bar{\mathbf{1}} = [1, \dots, 1]^T$ 로 정의된다.

<106> 단계 311은 분리필터 계수를 갱신하는 과정이다.

【수학식 24】

$$w_{ij}(f,b+1) = w_{ij}(f,b) + \mu \Delta w_{ij}(f,b)$$

<108> 상기 식(21)의 시간영역의 제약조건은 식(24)의 분리필터 계수를 계산한 이 후 분리필터 계수를 시간영역으로 변환하여 제약조건을 적용하는 것과 동일하다.

<109> 단계 312는 분리필터 계수의 크기를 주파수 영역에서 대각필터 계수의 전력 (norm)으로 나누어줌으로써 정규화하는 과정이다. 분리필터 계수의 정규화는 갱신 이 진행됨에 따라 분리신호의 전력이 무한히 작아지는 현상을 방지하는데 목적이 있다. 분리필터 계수의 정규화는 블록 시간마다 수행할 필요는 없으며 일정한 간격 으로 수행함으로써 계산량의 증가를 줄일 수 있다.

【수학식 25】

$$w_{ij}(f,b) = \frac{w_{ij}(f,b)}{\sqrt{\sum_{\text{for all } \alpha} |w_{ij}(\alpha,b)|^2}}$$

<111> 단계 313은 분리필터 계수가 원하는 상태로 수렴했는지를 판단하는 과정이다. 이 과정은 가용한 다채널의 혼합신호 샘플들을 모두 처리했거나 혹은 반 복해서 처리해야하는 회수만큼 모두 반복 수행한 경우 수렴으로 간주된다. 그 밖에 분리필터 계수나 그 밖의 변수들이 특정한 값으로 수렴하는 것을 조사하여 수렴을 판단할 수 도 있다.

<112> 단계 314는 수렴이 완료된 경우 가용한 다채널 혼합신호를 분리하여 최종 분 리신호를 얻는 과정이다. 이 과정은 식 (11)을 이용하며 역시 오버랩-세이브 방식

을 이용하여 주파수 영역에서 수행될 수 있다.

<113> 도 4a는 실제 사무실 환경에서 녹음된 음성과 음악을 두개의 마이크로 녹음한 다음 등을 이용하여 본 발명의 방법과 장치들을 이용하여 분리한 결과이다. 도 4의 파형은 위로부터 혼합신호 x_1 과 x_2 분리신호 u_1 과 u_2 를 보여준다. 여기서 $L=128, N=8L, M=L, \mu=0.0025$ 를 사용하였다. 도 4b는 최종적으로 얻어진 분리 필터이다.

<114> 혼합신호 본 발명은 음성 인식기나 음성 부호화기의 전 처리기 등으로 활용됨으로써, 여러 신호가 혼재된 상황에서 음질의 왜곡을 최소화하면서 효과적으로 분리할 수 있다. 도 5는 음악과 음성 원 신호(s_1, s_2), 가상의 혼합회로를 통해 혼합된 신호(x_1, x_2), 본 발명의 분리 방법을 분리한 신호(u_1, u_2)의 파형을 비교한 것이다. 도 5는 본 발명의 방법과 장치를 이용하여 분리한 신호의 파형은 원 신호의 파형과 거의 동일하며 우수한 음질을 갖는다는 것을 보여준다.

<115> 이상에서 설명한 본 발명은 전술한 실시 예 및 첨부된 도면에 의해 한정되는 것이 아니고, 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 여러 가지 치환, 변형 및 변경이 가능하다는 것이 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 있어 명백할 것이다.

【발명의 효과】

<116>

상기한 바와 같이 본 발명은 자연접선기울기를 이용한 다채널 블라인드 신호 분리를 통해 다중경로를 통해 혼합된 다채널 혼합신호를 효과적으로 분리할 수 있다. 음악과 음성이 혼합된 경우나 여러 사람의 음성이 혼합된 경우 등에서 음성인식율과 통화품질이 현저하게 개선되는 효과가 있다.

【특허청구범위】

【청구항 1】

다중경로 다채널 혼합신호를 분리하는 분리 방법에 있어서,

(a)혼합신호로부터 현재 프레임을 형성하는 단계(302);

(b)상기 프레임을 분리하여 현재 프레임의 분리신호를 형성하는 단계(303);

(c)상기 분리신호를 비선형 함수로 변환하는 단계(304);

(d)상기 분리신호와 비선형 변환된 신호로부터 에일리어싱이 제거된 정규화된 상호전력 스펙트럼을 계산하는 단계(305);

(e)상기 에일리어싱이 제거된 정규화된 상호전력 스펙트럼을 이용하여 자연 접선기울기를 계산하는 단계(310);

(f)상기 자연접선기울기를 이용하여 분리필터 계수를 갱신하는 단계(311);

(g)상기 분리필터 계수를 정규화하는 단계(312);

(h)수렴을 판단하여 수렴할 때 까지 상기 (a)-(g) 단계들을 반복하는 단계(313); 및

(i)수렴이 완료된 이후 분리필터 계수를 이용하여 다채널 혼합신호를 분리하는 단계(314)

를 포함하는 것을 특징으로 하는 다중경로 혼합신호의 블라인드 분리를 위한 방법.

【청구항 2】

제 1항에서, 단계 (b)는

(b1)상기 분리필터 계수와 상기 현재 프레임의 혼합신호를 주파수 영역으로 변환하는 단계;

(b2)주파수 영역에서 상기 분리신호를 계산하고 그 결과를 시간 영역으로 변환하는 단계;

(b3)시간 영역으로 변환된 분리신호의 처음 L 샘플을 0으로 치환하는 단계;

를 포함하는 것을 특징으로 하는 다중경로 혼합신호의 블라인드 분리를 위한 방법.

【청구항 3】

제 1항에서, 단계 (d)

(d1)상기 현재 프레임의 분리신호와 비선형함수 변환신호를 주파수 영역으로 변환하는 단계;

(d2)상기 주파수 영역으로 변환된 분리신호와 비선형 함수 변환신호로부터 현재 프레임의 상호전력 스펙트럼을 계산하는 단계(306);

(d3)상기 주파수 영역으로 변환된 분리신호와 비선형 함수 변환신호의 평균 전력 스펙트럼을 계산하는 단계(307);

(d4)상기 상호전력 스펙트럼을 정규화하는 단계(308);

(d5)상기 정규화된 상호전력 스펙트럼을 시간 영역으로 변환한 다음 시간 영역에서 제약조건을 가하여 처음 L 샘플만을 취하고 다시 주파수 영역으로 변환하는 단계(309);

를 포함하는 것을 특징으로 하는 다중경로 혼합신호의 블라인드 분리를 위한 방법.

【청구항 4】

제 1항에서, 단계 (e)는

(e1)상기 에일리어싱이 제거된 정규화된 상호전력 스펙트럼의 대각원소에 대한 기댓값 1을 제거하여 논홀로노믹(nonholonomic) 조건을 만족시켜주는 단계;

(e2)정규화된 상호전력 스펙트럼의 대각원소를 계산한 이후 주파수 영역의 분리필터 계수와 곱하여 자연접선기울기를 계산하는 단계;

를 포함하는 것을 특징으로 하는 다중경로 혼합신호의 블라인드 분리를 위한 방법.

【청구항 5】

제 1항에서, 단계 (f)는 상기 자연접선기울기를 이용하여 최대기울기 방법을 통해 분리필터 계수를 갱신하는 것을 특징으로 하는 다중경로 혼합신호의 블라인드 분리를 위한 방법.

【청구항 6】

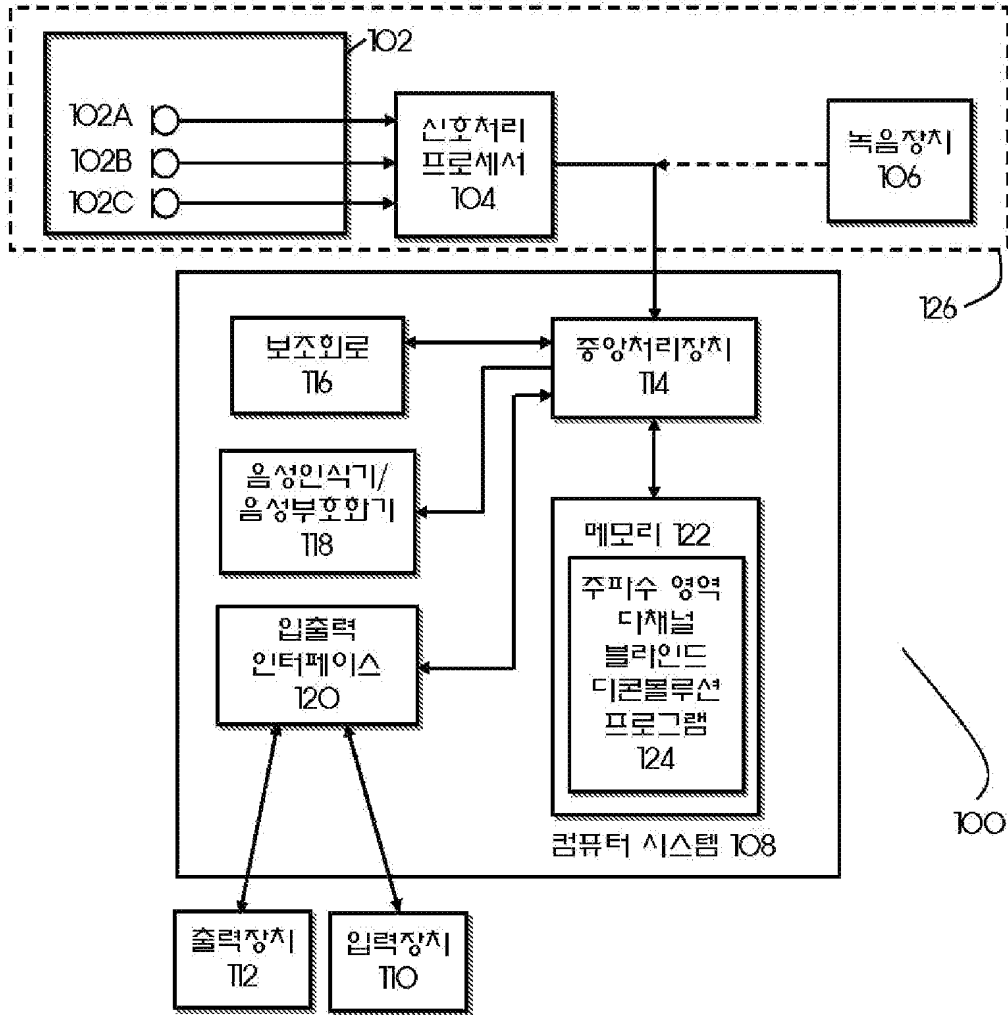
음성인식기나 음성부호화기가 결합되는 것이 특징인 다중경로 혼합신호의 블라인드 분리를 위한 장치.

【청구항 7】

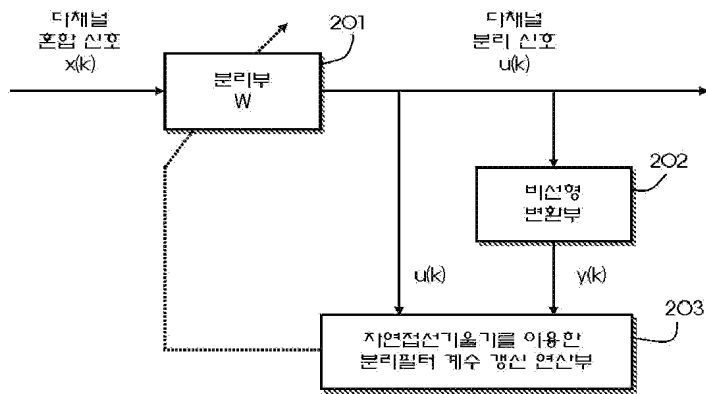
상기 1항의 다중경로 다채널 혼합신호의 블라인드 분리 방법을 내장하고 있는 컴퓨터 저장매체.

【도면】

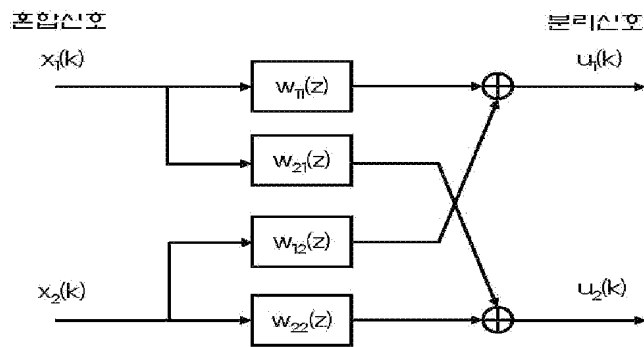
【도 1】



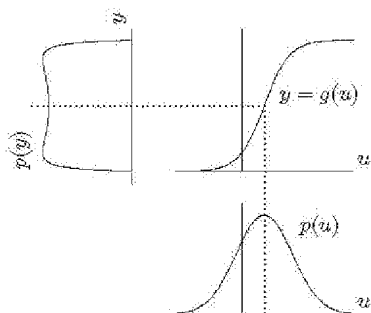
【도 2a】



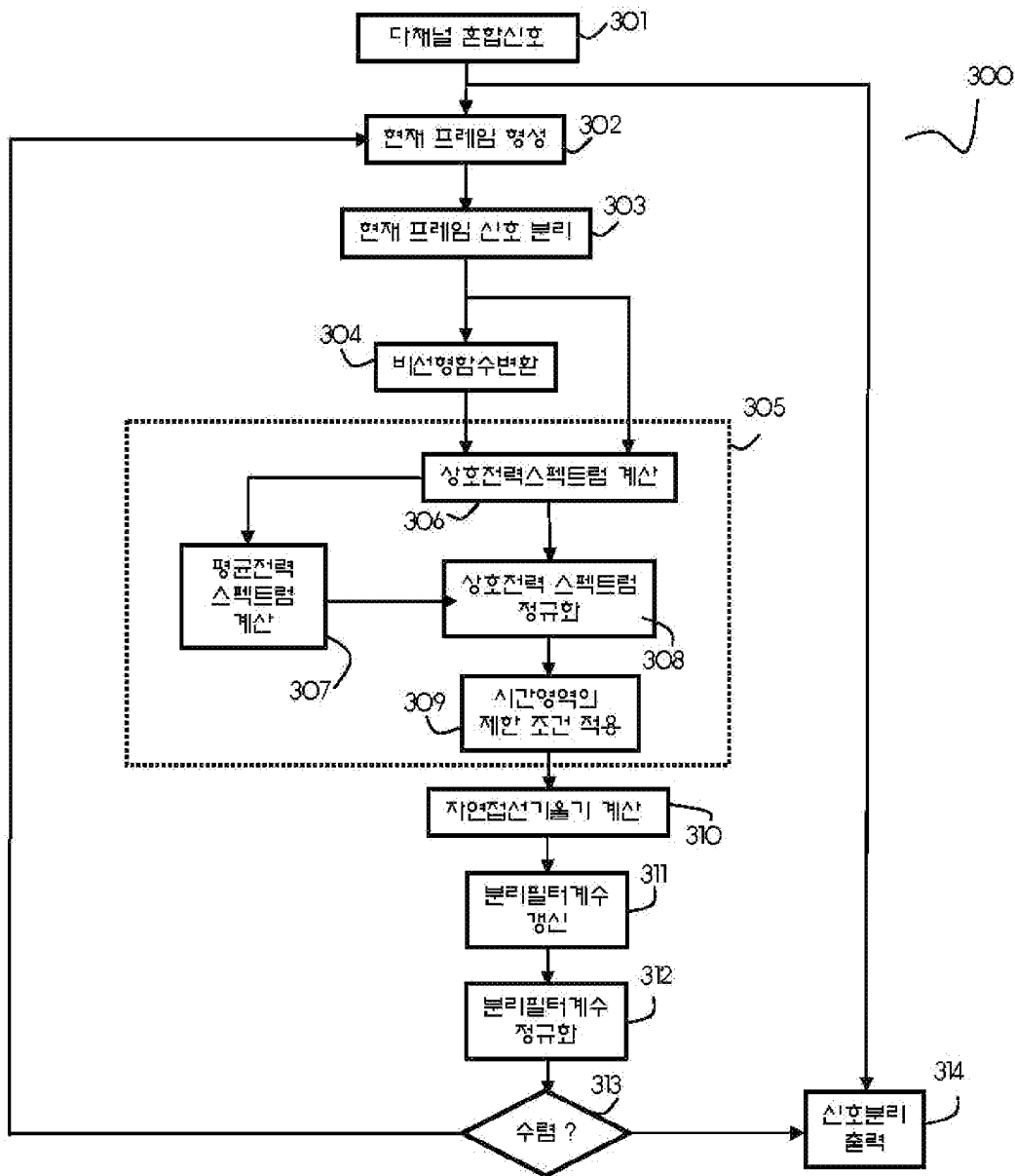
【도 2b】



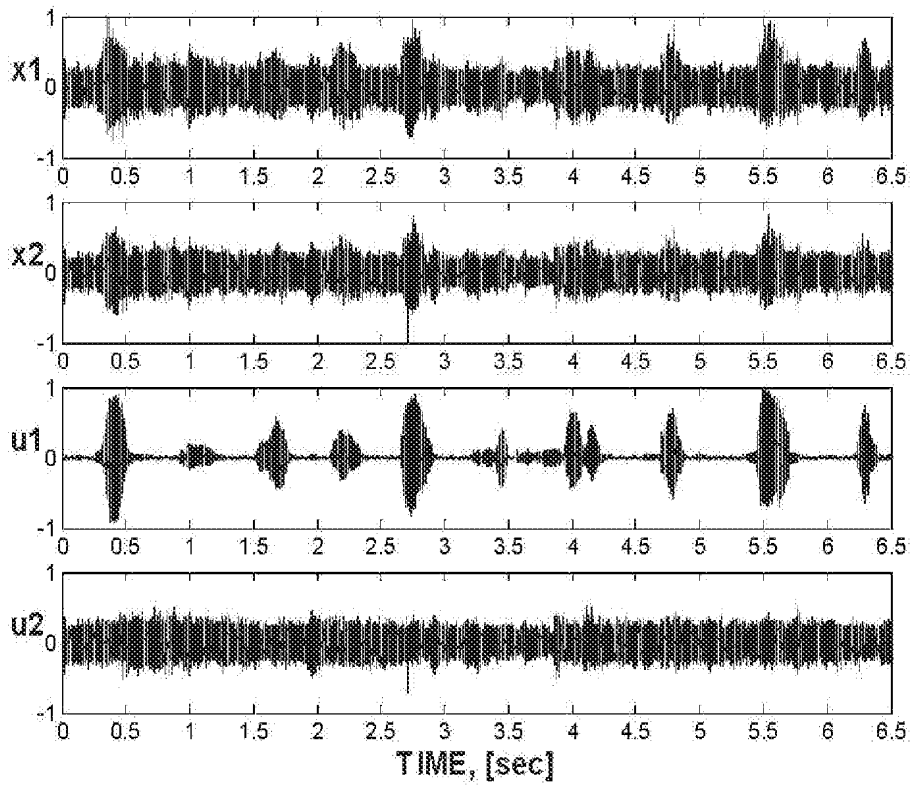
【도 2c】



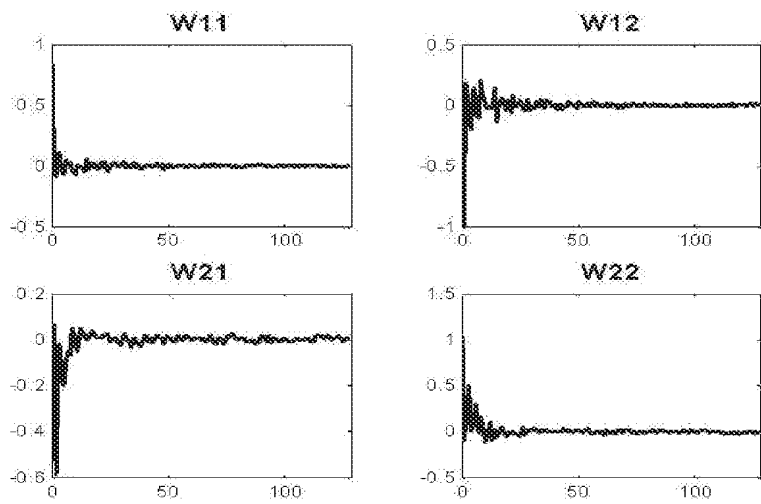
【도 3】



【도 4a】



【도 4b】



【도 5】

